# Chapter 4

# The Effectiveness of GIS In Secondary Education: Experiments In Three High Schools

## Introduction

This dissertation analyzes the implementation and effectiveness of GIS in secondary education. Analysis of the survey results in Chapter 3 provided an understanding of the *implementation* of GIS in a national context. To supplement the national assessment with a detailed assessment at the local level, this chapter analyzes the *effectiveness* of GIS within three high schools. In order to evaluate GIS within individual high schools, a series of experiments (Chapter 4) and case studies (Chapter 5) were conducted.

The literature review revealed that although GIS may improve self-confidence, ability to work in groups, and skills in integrating large amounts of information, there is little evidence that these same students could better understand the basic skills and content of the subject matter than if they used traditional materials. One teacher responding to the national GIS education survey, although experienced with GIS, expressed concerns about its effect on learning:

> "I personally have been troubled with the question of whether students are learning geographic inquiry strategies or merely learning to use a very powerful tool without much thinking about the underlying questions under consideration."

Another respondent wrote, “I’m not convinced that this [GIS] is worth the effort, but the potential is there.”

Most of the research on educational GIS is comprised of anecdotal accounts of implementation in individual classrooms, rather than assessing the effectiveness of GIS through experiments. Assessing the difference that GIS makes in teaching and learning in the classroom requires experiments. “Of all empirical work, comparative experiments provide the strongest evidence about the effects of education interventions” (Porter 1997: 523). The experiments conducted as part of this dissertation research sought to supplement anecdotal accounts with numeric data. Another goal was to minimize external variables by conducting the experiments with the same lessons, with the same software, and with students in the same grade levels.

To be able to make informed decisions about implementing and evaluating a GIS, qualitative and quantitative data are needed. Herman (1994) advocated the use of both quantitative and qualitative methods, and a variety of indicators to document effects. The goals of both the case studies and experiments were to analyze the catalysts and challenges to the use of GIS technology in the secondary curriculum, and to determine if GIS technology has an effect on the teaching method and student acquisition of geographic content and skills.

## Research Methodology

### *Selection of Experimental and Case Study Schools*

The experiments were conducted in three high schools in metropolitan Denver, Colorado, USA. These schools were Riparian High School, Hope High

School, and Prairie Vista High School[1]. All are public four-year high schools enrolling between 1,200 to 3,000 students. The schools were selected based on criteria aimed at ensuring that the schools, courses, teachers, and students would be as equivalent as possible so that experimental results could be compared. First, each school had to have an active geography program, where geography is taught as a separate, distinct subject, rather than as part of science, government, or history. Geography needed to be taught in at least one class for at least one full academic year. The same teacher must have taught geography in the school for at least two full years prior to the experiments.

The teacher in each school needed to be actively using geographic information systems software in the curriculum. This requirement ensured that GIS-based lessons could be implemented, observed, and assessed. All schools had to use the same GIS software—*ArcView,* version 3.1. The teachers must have previously used or were currently using national or state geography content standards in their geography instruction, so that the geography standards could be used to assess the GIS-based lessons. Each teacher had to be willing to work with the author in implementing the technology and the experiments in their classrooms.

### ***Experimental Design***

#### *Description*

The experimental design included the creation of 12 geography lessons, each with two versions—a GIS-based version, and a version using traditional print materials—textbooks, paper maps, atlases, and data tables. To avoid mixing students between control and experimental groups, whole sections, or class periods,

[1] The names of the schools have been changed to protect the anonymity and confidentiality of students and teachers.

of students were kept in the same group, rather than splitting a class in half. In each of the three schools, experimental groups were comprised of sections, or class periods, of a geography course in which students used GIS to complete the lessons. Control groups were comprised of other sections in the same geography course in which students were given traditional print materials to complete the lessons. At no time during the semester did the control group students use GIS.

In other words, for a certain geography course, students in Periods 1 and 2 may have been assigned to the experimental group, while students in Period 4 may have been assigned to the control group. The one exception was in an advanced geography class in Riparian High School that contained only one section that could not be split apart. In this case, students from a previous year who did not use GIS were used as the control group, and students from the current year were used as the experimental group.

To minimize instructional variation, the number of teachers teaching the lessons was kept small—one in Riparian, one in Prairie Vista, and two in Hope High School. Although the characteristics of the teacher, the teacher's experience with GIS, the school curriculum, computer technology and access, classroom structure, and student backgrounds varied among the three schools, the experimental design attempted to minimize as many differences as possible. If all other variables had been reasonably controlled for, differences between pre- and posttest results could be attributed to the use of GIS technology and methods. I acted as participant observer in the experiments, taking notes for the case studies and offering assistance to both the control and the experimental groups.

Experiments were conducted during Fall 1998 at Hope High School and during Spring 1999 in all schools, with students in Grades 9, 11, and 12. Each experimental group completed at least two of the same lessons as the control group during each semester. All except two of the lessons were taught in all three schools with the same grade levels, to allow comparisons to be made across schools.

*Pretest and Posttest Design*

A pretest and a posttest were given to each student at the beginning and the end of each semester in which the experiments were conducted. Pretests and posttests consisted of (1) “standardized tests”—based on national, state, and district geography standards, and (2) a spatial analysis test that I created. Two standardized tests were given. One was a competency-based geography test created by the National Council for Geographic Education (1983), given in Riparian and Prairie Vista high schools (Appendix A.7). The test contains 75 four-option, multiple-choice items on physical and human geography. The teachers selected 53 of the four-option, multiple-choice items for Grade 9 students and 72 for Grade 11 and Grade 12 students in a multiple choice format. Students answered questions based on the small maps, stories, and charts included in the test. Teachers selected questions that most closely aligned with the school district’s K-8 curriculum, and would therefore cover topics that students were most familiar with. Also, questions were selected that did not refer to a place that had changed names or otherwise give an indication that the test was over 10 years old. Questions on both physical and human geography were included. The other standardized test, given in Hope High School, was created by the County Assessment Board and based on national geography standards 1 (using maps) and 4 (physical and human characteristics of places)(Appendix A.8). It required students to fill in answers to a narrative of a trip around the world using *Goodes World Atlas.* The maximum standardized score was 100.

The spatial analysis test required students to choose the three best sites in a community for a “Spiffy’s” fast food restaurant, to spot the sites on a map, to rank the factors used in choosing the sites, and to support each choice with a short essay (Appendix A.9). Students were given printed maps of streets showing traffic volume, existing Spiffy’s restaurants, locations of high schools, percent of population aged 15 to 24 by census tract, annual median income, and zoning. Students also were asked to rank these six variables from 1 to 6, with “1” considered as most important to site the restaurant. They were given 15 minutes to complete the test.

Because the spatial analysis test was designed as an alternative, problem-solving activity, alternative assessments were used to score it. The frequency of citing geographic variables, completeness, clarity, and evidence of content knowledge of geography were equally important. The minimum score possible was 4; the maximum score was 20 (Table 4.1).

Although no single correct answer existed, several sites clearly would reveal errors in understanding geographical concepts, such as siting the establishment in the middle of an intersection, in an area zoned residential or industrial, or adjacent to an existing Spiffy’s restaurant. The assessment’s overall goal was to measure whether the student could use data in spatial analysis, and build a good case for his or her decision.

Table 4.1. Scoring Guide for Spatial Analysis Test.

| Component | Score | | | | |
|---|---|---|---|---|---|
| | **Unsatisfactory = 1** | **Minimal = 2** | **Rudimentary = 3** | **Commendable = 4** | **Exceptional = 5** |
| **Frequency of Citing Geographic Variables** | Does not cite any variables. | Cites 1 variable. | Cites 2 variables. | Cites 3 variables. | Cites 4 or more variables. |
| **Complete-ness; Attention to Detail** | No detail provided. | Minimal detail provided: One sentence. | At least two sentences of meaningful detail provided. | At least three sentences of detail provided. | At least four detailed sentences provided. |
| **Clarity** | Does not provide any explanation or explanation is incoherent. | Provides little explanation or explanation is difficult to interpret. | Provides adequate explanation that can be understood. | Locations and explanations are thorough. | Locations and explanations are as clear as possible. |
| **Evidence of Content Knowledge** | Provides no evidence of content knowledge and spatial analysis is completely in error. | Provides little evidence and spatial analysis contains at least 2 errors. | Provides some evidence but spatial analysis contains at least 1 error. | Demonstrates knowledge of at least one spatial relationship and spatial analysis is sound. | Demonstrates knowledge of at least two spatial relationships and spatial analysis is thorough and correct. |

I designed and assessed lessons consulting Daugherty's (1992) dimensions of relevant assessment, using a performance, or alternative, assessment model. This model emphasizes the production of something to demonstrate knowledge and skill development and progress, rather than selecting from a multiple-choice exam (Okey 1995). These lessons require students to accomplish tasks to solve authentic problems. Following the recommendations by Herman et al. (1992), tasks involve portfolios, exhibits, investigations, demonstrations, written or oral responses, and journals. Performance-based assessment should be used where a student demonstrates that he or she can create an answer or a product that demonstrates his or her knowledge or skills (Rudner and Boston 1994). The "dimensions of learning" model asserts that when process skills and knowledge are complex, performance assessment is a must (Marzano et al. 1993). Through these lessons, the teachers and I attempted to engage students in worthwhile, significant, and meaningful tasks, and evaluate them in an authentic assessment (following Hart 1994). Authentic assessment means that the student is evaluated *within* the context of what they are doing, not just afterward.

I assessed all tests and lessons except for the standardized tests, which were scored by the teachers. To maximize internal validity—the degree to which the correlation between independent and dependent variables was caused by the independent variable—the tests were assessed in a random order, to avoid biasing any one group with higher or lower scores. Evaluations for all lessons were based on the geographic skills identified in the national geography standards—asking geographic questions, acquiring geographic information, organizing geographic information, analyzing geographic information, and answering geographic questions. Scoring guides were created in an attempt to measure degrees of quality, quantity, frequency, and understanding to capture elements of critical thinking and performance.

Because I did not have complete control over the experiments, they should be considered "quasi-experiments." I did not solely conduct all the lessons nor was I able to dictate which students would be in the control and experimental groups. Groups were assigned randomly as part of regular high school course and room assignments, with a random mix of male and female students in each class. Data on gender were obtained from the case studies, from the student's name, or from the teachers.

The experimental design's goal of internal validity was met except for the criterion that students were not selected based on ability. External validity—the degree to which these results are applicable and generalizable to other groups—was attempted by making the experiments as equivalent as possible among the schools, students, and classes. Any individual within the school had an equal chance of being selected for the study, but it was not a true random sample statistically. These high school students were considered typical American high school students—from averaged-sized, urban and suburban schools, with average class sizes. Precision—that the differences in the dependent variable really are detectable and attributable to the independent variable—was addressed through statistical measures and by conducting the experiments in not one but three different high schools.

*Statistical Procedures*

One cannot separate the effects of technology from the quality of the instruction and curriculum in which it is embedded (Herman 1994: 151). It is not technology in and of itself that makes a difference, but the way it is used. The national survey indicated that the use of GIS *technology* appears to change the teaching *method* from a traditional teacher-centered approach to an exploratory, constructivist, student-centered approach. Therefore, GIS technology and methods were treated as one independent variable for the experiments. GIS was tested for its influence on the dependent variables—knowledge of geography content and geographic skills. Content and skills were defined by scoring guides based directly on the national geography standards (Geography Education Standards Project 1994) or indirectly through state of Colorado and school district standards that were based on the national standards.

Two-sample *t*-tests with equal variances were conducted on both the standardized and the spatial analysis tests, to discover any significant difference between control and experimental groups when the experiments began. Posttest scores were compared to determine if GIS caused any significant difference in student's geographic knowledge and skills. Pretest scores were also compared to posttest scores via paired *t*-tests to determine the amount of change over the semester, and whether the change was significantly different between each group. These provided data about whether students using GIS may have learned a different amount of content and skills than students using traditional methods and materials. Scores from the lesson modules were also analyzed with a two-sample *t*-test. A one-way analysis of variance (ANOVA) was conducted for all sections of Grade 9 geography at Riparian High School to determine if GIS affects overall achievement and achievement on the tests used in the experiments. ANOVAs and *t*-tests provided data on gender differences. Several regression models were established to investigate the relationship between GIS, pretest scores, and the difference between pretest and posttest scores. These linear and non-linear models helped determine the effect of independent variables on student learning. Analyses were conducted on individual classes, between classes, in individual schools, and grouping all schools together to determine if any of the findings could be generalized to more than one class or school. The research sought to identify cause and effect in an attempt to isolate the effect of technology and methods.

*Interpreting the Data*

Experiments are performed to test hypotheses and to establish causality; that is, do the independent variables precede the dependent variables in time? It must be determined if the differences were not the results of chance, as established by the rules of probability.

Considering Group C as the control group and Group E as the Experimental Group in each high school, let: $C_0$ = Group C pretest results, $E_o$ = Group E pretest results, $C_1$ = Group C posttest results, and $E_1$ = Group E posttest results. If $C_0$ is not equal to $E_o$, a difference exists between the starting characteristics of the groups. If the difference is significant, then the groups were unequal to begin with and it will be difficult to attribute subsequent differences of the effect of GIS. If the groups meet the criteria of $C_0 = E_0$, no significant difference between starting characteristics of groups existed, and the following decision matrix guided the interpretation of the data. Symbols > and < indicate *significantly* higher and lower, respectively (Table 4.2). The same decision matrix was used in all three high schools.

Table 4.2. Decision Matrix for Interpreting Test Scores in Experiments.

| **Decision Matrix** | | | |
|---|---|---|---|
| **Control Posttest vs. Pretest, Experimental Posttest Vs Pretest** | **Control Posttest vs. Experimental Posttest** | **Effect of GIS** | **Explanation** |
| $C_1>C_0$ and $E_1>E_0$: Scores improved from Pre- to Post-test for both groups. | $C_1>E_1$ | Positive | Scores improved for both groups. GIS modules were effective, but not as effective as traditional methods. |
| | $C_1=E_1$ | Positive | Scores improved for both groups. GIS and traditional methods were equally effective. |
| | $C_1<E_1$ | Positive | Scores improved for both groups. GIS methods were even more effective than traditional methods. |
| $C_1=C_0$ and $E_1=E_0$: Posttest scores same as pretest for both groups. | $C_1>E_1$<br>$C_1=E_1$<br>$C_1<E_1$ | No effect | Scores did not improve for either group. GIS and traditional methods were ineffective in terms of scores. |
| $C_1<C_0$ and $E_1<E_0$: Scores decreased from Pre- to Posttest for both groups. | $C_1>E_1$ | Negative | Scores decreased for both groups. Scores decreased even more for group using GIS than group using traditional methods. |
| | $C_1=E_1$ | Negative | Scores decreased equally for both the control and experimental groups. |
| | $C_1<E_1$ | Negative | Scores decreased for both groups. Scores decreased more for the group using traditional methods than GIS methods. |

## Analysis of Experiments at Riparian High School

### *Description of Geography Program and GIS Implementation*

The geography curriculum at the high school consists of *Geography 1*, a year-long required class for Grade 9 students, and *Geography 2*, a semester-long elective class for Grade 11 and 12 students. Mr. Warren E. Stevenson[2] has been teaching geography for 29 years, including 9 years at this high school and 20 years in a middle school in the same community. The implementation of GIS at Riparian High School began in 1996 as an initiative between Mr. Stevenson and this author.

### *Description of Experiments*

Experiments conducted at Riparian High School took place during Spring semester 1999 (Table 4.3). Mr. Stevenson was the teacher for all lessons with some assistance for each class from a student teacher and student intern.

Experiments were conducted as planned except that the semester's schedule prohibited *Geography 2* students from taking the standardized test at the end of the semester. Laboratory difficulties, which will be described in the case study, did not allow all students to complete the *Earthquakes Everyday* GIS-based lesson. Africa Unit 3 was dropped from the analysis because experimental group students were not able to complete it.

[2] The names of the teachers and students were changed throughout this document to protect their anonymity and confidentiality.

Table 4.3. Experiments Conducted in Riparian High School.

| **Experi-ment** | **Class** | **Grade Level** | **Lessons** | **Description** | **Pretest/ Posttest** |
|---|---|---|---|---|---|
| **1:**<br><br>Compare C1 Vs. E1 | *Geo-graphy 1* | 9 | Earthquakes Everyday<br><br>Africa Regional Analysis Units 1 and 2 | Control Group = Period 1 (730-830am) and Period 3 (930-1030am)<br><br>Experimental Group = Period 5 (12-1pm) and Period 7 (2-3pm) | Standardized pretest and posttest<br><br>Spatial Analysis pretest and posttest |
| **2:**<br><br>Compare C2 Vs. E2 | *Geo-graphy 2* | 11 and 12 | *The Hill* Neighbor-hood Analysis | Control Group = Spring 1996 students<br><br>Experimental Group = 1999 Period 4 (1030-1130am) | Spiffy's Spatial Analysis Pretest / Posttest for Experimental Group only |

In Experiment 1, three GIS-based lessons were given to the two experimental groups, while the same lessons were given to the two control groups using atlases and a textbook. In the *Earthquakes Everyday* lesson, the control groups used the class textbook, a preprinted list of the previous week's earthquake locations by latitude and longitude, and printed atlases. Students plotted earthquake locations using pencils and a printed map, and used the atlases for the additional data such as plate boundaries and cities. The same earthquake lesson was taught to the experimental groups using the Internet and GIS. Students used *ArcView* GIS to plot and analyze locations and patterns of earthquakes. The Africa lessons were taught in a similar way—with traditional materials for the control group, and GIS for the experimental group. In Experiment 2, experimental group students completed *The*

*Hill* lesson using GIS. All experimental group students worked through an *ArcView* tutorial before beginning the lessons. Each group received identical questions.

### *Tests for Group Differences*

A two-sample *t*-test with equal variances was run on both types of pretests (standardized and spatial analysis), with one group as the control group, and one group as the experimental group. Neither the spatial analysis *t*-test nor the standardized *t*-test showed a significant difference between the control and experimental groups (Table 4.4).

Table 4.4. Tests for Group Differences, Riparian High School.

| Experi-ment | Class | n | Mean | Standard Deviation | Difference | t, P values |
|---|---|---|---|---|---|---|
| Spatial Analysis Test | Control | 51 | 12.9412 | 4.0714 | .3445 df=119 | t= 0.4552 P=0.6499 |
| | Experimental | 70 | 13.2857 | 4.1394 | | |
| Standar-dized Test | Control | 53 | 61.9082 | 18.8790 | 1.3389 df=120 | t= 0.4613 P=0.6454 |
| | Experimental | 69 | 63.2470 | 13.1546 | | |

Therefore, we can have confidence with subsequent tests to compare the group using GIS and the group using traditional methods and materials, since there is a good chance that both groups are drawn from the same population.

### *Assessing Spatial Analysis Tests*

To assess the difference between pretest and posttest scores, paired *t*-tests were conducted for each class. Paired *t*-tests were conducted for both the spatial analysis tests and for the standardized tests. Spatial analysis scores for one control and two experimental groups *declined* significantly (Table 4.5).

Table 4.5. Results of Paired *t*-tests, Spatial Analysis Test, Riparian High School (*=significant at ∀=.05)

| Class Period | Test | n | Mean | Standard Deviation | Difference | t, P values |
|---|---|---|---|---|---|---|
| 1 Control | Pretest | 23 | 12.2174 | 4.4206 | -1.3478 | t=- 1.1449 P=0.2645 |
| | Posttest | 23 | 10.8696 | 4.9203 | | |
| 3 Control | Pretest | 25 | 13.9200 | 3.7519 | -6.8000 | t= -6.8114 P=0.0000* |
| | Posttest | 25 | 7.1200 | 4.6755 | | |
| 4 Experi-mental | Pretest | 9 | 13.7778 | 3.7342 | -0.8889 | t= -0.4516 P=0.6635 |
| | Posttest | 9 | 12.8889 | 3.8873 | | |
| 5 Experi-mental | Pretest | 23 | 12.8696 | 3.6220 | -3.2174 | t= -2.7395 P=0.0120* |
| | Posttest | 23 | 9.6522 | 4.4885 | | |
| 7 Experi-mental | Pretest | 20 | 14.8500 | 3.3289 | -2.1500 | t=-2.5202 P=0.0208* |
| | Posttest | 20 | 12.7000 | 4.4851 | | |

A paired *t*-test was conducted for all control group classes in Riparian High School, comparing spatial analysis pretests and posttests to find out if students truly performed worse on the posttests overall. Scores declined significantly for all control groups as a whole (Table 4.6). Scores declined significantly for the experimental group as well, though not as much as for the control group (experimental t=3.44 vs. control t=4.89).

Table 4.6. Results of Paired *t*-tests for Control and Experimental Groups as a Whole, Spatial Analysis Test, Riparian High School (*=significant at ∀=.05)

| Group | Test | n | Mean | Standard Deviation | Differ-ence | t, P values |
|---|---|---|---|---|---|---|
| All Control Groups | Pretest | 48 | 13.1042 | 4.1321 | -4.1875 | t=-4.8886 P=0.0000* |
| | Posttest | 48 | 8.917 | 5.1067 | | |
| All Experi-mental Groups | Pretest | 52 | 13.7885 | 3.5773 | -2.4038 | t=-3.4460 P=0.0011* |
| | Posttest | 52 | 11.3846 | 4.5811 | | |

In all cases, posttest standard deviation was greater than the pretest, indicating that while some students showed an improvement in their spatial analysis, others did just as poorly or worse than they did at the time of the pretest.

The 1988 National Assessment of Educational Progress test, given to 3,000 Grade 12 students, indicated that most are not well-versed in geography (Bettis 1997). Thus, low spatial analysis test scores should come as no surprise. Students may have done worse on the spatial analysis test at the end of the semester for several reasons. At the end of the semester, they may have recognized that their final grades would not be affected by their performance on the test and thus had a disincentive to do well. Also, the shootings at a nearby high school during the previous month caused apathy and psychological trauma among the entire high school community. Indeed, many students scored a 4 or below out of 20 because they did not write any explanation for their site selections. However, a *t*-test run only on scores above 4 resulted in a similar *t*-statistic as for the entire school sample ($t=2.97$; $P=0.006$ for control group; $t=3.01$; $p=0.005$ for experimental group). Scores still declined significantly, even for students who had filled out the whole exam. As measured by these tests, GIS did not appear to improve students' performance on the spatial analysis test.

### *Assessing Standardized Tests*

*T*-tests on *standardized* tests were run to assess the difference that GIS made on student learning as measured by pretest versus posttest scores (Table 4.7). Opposite to the spatial analysis trend, students significantly improved their performance on the standardized test in each class, slightly more if they used GIS.

Table 4.7. Results of Paired *t*-tests on Standardized Tests, Riparian High School (*=significant at ∀=.05)

| Class Period | Test | n | Mean | Standard Deviation | Difference | t, P values |
|---|---|---|---|---|---|---|
| 1 Control | Pretest | 24 | 62.6572 | 19.2012 | 9.6698 | t=3.9990 P=0.0006* |
| | Posttest | 24 | 72.3274 | 13.7041 | | |
| 3 Control | Pretest | 24 | 64.2296 | 14.5352 | 5.1887 | t=2.3522 P=0.0276* |
| | Posttest | 24 | 69.4182 | 12.5647 | | |
| 5 Experi-mental | Pretest | 24 | 64.0723 | 14.5777 | 5.5818 | t=2.4407 P=0.0228* |
| | Posttest | 24 | 69.6541 | 14.7826 | | |
| 7 Experi-mental | Pretest | 22 | 61.6638 | 10.1086 | 8.4906 | t=4.2274 P=0.0004* |
| | Posttest | 22 | 70.1544 | 10.8409 | | |

Combining control and experimental groups yielded the same results (Table 4.8).

Table 4.8. Results of Paired *t*-tests on Standardized Tests for Control and Experimental Groups as a Whole, Riparian High School (*=significant at ∀=.05).

| Group | Test | n | Mean | Standard Deviation | Difference in Means | t, P values |
|---|---|---|---|---|---|---|
| All Control Groups | Pretest | 48 | 63.4434 | 16.6853 | 7.4292 | t=4.4980 P=0.0000* |
| | Posttest | 48 | 70.8726 | 13.0889 | | |
| All Experi-mental Groups | Pretest | 46 | 62.9204 | 12.5611 | 6.9729 | t=4.5565 P=0.0000* |
| | Posttest | 46 | 69.8934 | 12.9074 | | |

Similar to the findings from the spatial analysis test, no significant difference between control and experimental groups was noted. The hypothesis that GIS made a significant difference must be rejected using either test.

### *Item-By-Item Analysis of Standardized Test*

The standardized test was a measure of basic geographic content and skills as created by the NCGE, rather than a measure of the content and skills covered in the three GIS-based lessons that the students completed during the semester (*Earthquakes Everyday* and *Africa 1* and *2*). Therefore, to more directly link assessment to content, the standardized test was analyzed to determine if any items directly measured the content taught in the GIS-based lessons. Eight items on the standardized test were covered in these lessons. The number of incorrect answers to each of these items was calculated to assess if GIS made a difference in the item scores (Table 4.9).

Table 4.9. Item-By-Item Tally, Standardized Test, Grade 9, Riparian High School.

| | Standardized Test Item Number | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Group** | **3** | **4** | **9** | **16** | **21** | **26** | **27** | **48** |
| Experimental (GIS), Number Missed | 19 | 23 | 1 | 19 | 11 | 4 | 28 | 19 |
| Control (No GIS), Number Missed | 23 | 16 | 1 | 16 | 10 | 3 | 19 | 25 |

The control group did better on five items, the experimental group did better on two items, and the two groups tied on one item. The control group scored incorrectly 113 times on these eight questions, while the experimental group missed these questions 124 times. It does not appear from the item-by-item analysis that GIS made a difference in content and skills measured by these tests.

Up to this point, student content and skills have been assessed using a pretest/posttest design before and after lessons were used during the semester. These lessons were identical, except that the control group used traditional materials, such as books and printed maps, and the experimental group used

*ArcView* GIS. A comparison of student performance *on the actual lessons* would be instructive to assess the difference whether GIS made any difference, and how much the difference was. Performance assessments are potentially more valid because the tasks themselves are direct measurements of learning goals.

### ***Assessing Africa Regional Geography Lessons***

The Africa units (Appendices A.10-A.23) were originally selected because students had demonstrated less knowledge about Africa than any other continent. The lessons included analyzing surface features and physical characteristics, population and cultural characteristics, human and natural resources, and natural hazards. Each Africa module required two to three days to complete.

A two-sample *t*-test conducted on scores from the first Africa module showed a significant difference between the control and experimental group at the .10 level ($t=-1.7829$; $P=0.0778$). The experimental group mean was 1.18 point higher than for the control group (15.75 vs. 14.57). A two-sample *t*-test conducted on scores from the second Africa module showed a significant improvement from the control to the experimental group, this time at the .05 level ($t=-3.1327$; $P=0.0023$). For both lessons, students using GIS demonstrated significantly improved knowledge and skills.

### ***Assessing Earthquakes Everyday Lesson***

In the *Earthquakes Everyday* unit, students analyze the spatial pattern of earthquake data from the past week, and compare that pattern to plate boundaries, fault lines, cities, and gross domestic product. However, an analysis of the scores showed that the control group's mean score was 83.19 (out of 100) while that of the experimental group was only 38.97. The respective standard deviations also differed substantially—13.31 and 36.30, respectively. A histogram of scores and a review of student work revealed the reason why—most of the students using GIS were not able to complete the lesson, and left most of the items blank. The reasons the GIS students did not finish the test will be examined as part of the case studies in Chapter 5. The lowest score for the control group was 56, but 22 students in the experimental group scored less than 56, and 9 scored a zero. Because of this situation, I judged the test to be invalid for comparison purposes.

However, I determined that analyzing one question on the test would be valid. After plotting the earthquake epicenters by hand (control group) or importing them into the GIS (experimental group), students were asked "what are three noticeable characteristics of your pattern of earthquakes?" This question was selected because it best captured the objective of the lesson—to examine global patterns of earthquake occurrence. The question occurred early in both versions of the lesson, so nearly every student answered it. The question was scored by the following method (Table 4.10):

Table 4.10. Scoring Guide for *Earthquakes Everyday* Item, Maximum Score = 12.

| Assessment Criteria | Rudimentary 1 point | Acceptable 2 points | Superior 3 points |
|---|---|---|---|
| **Evidence of Spatial Analysis** | Little or no analysis | Some analysis | Shows interconnection among characteristics |
| **Earthquake Pattern Characteristic 1** | Little or no description | Some description | Very descriptive |
| **Earthquake Pattern Characteristic 2** | Little or no description | Some description | Very descriptive |
| **Earthquake Pattern Characteristic 3** | Little or no description | Some description | Very descriptive |

The mean score for the control group was 7.58, and the mean score for the experimental group was 6.37. No significant difference was detected between control and experimental groups (t=1.3766; P=0.1725); indeed, the students using GIS scored a bit lower. GIS did not make a difference in scores as measured by this section of the *Earthquakes Everyday* lesson.

### ***Assessing The Hill Neighborhood Analysis Lesson***

In a two-week unit, *Geography 2* students examined census housing and demographic data to compare *The Hill* neighborhood in the city to other neighborhoods in the community (Appendix A.6). In a research study reported in the *Journal of Geography*, Mason (1972) claimed that *The Hill* neighborhood fit the ghetto model based on five criteria—spatial enclave, minority status, social disorganization, inferior status, and lack of choice. He defined its spatial zones to be the regional core, transition zone, fringe, and "beyond the fringe." Students are first asked to read a modified, slightly simpler version of the article. The reading usually comes as a surprise to students, who do not expect to find these characteristics applied to a neighborhood adjacent to their high school, and in a city that is nationally perceived as affluent.

Students are then asked to analyze demographic characteristics and discuss whether they believe the author's claims about *The Hill* are valid. Experimental group students were required to select and map 12 to 16 characteristics from a table of approximately 60 variables to support their position. These variables, by census tract and block group, included age, ethnicity, median income, median rent, housing tenure (owning versus renting), and number of units per dwelling. Students were assessed on their sets of maps and narrative reports that defend their position on the issue.

Because there was only one *Geography 2* class, it was not feasible to break up the class into a control and experimental group. Even had the space and instructor existed to do so, the students would have consulted with each other, as the class is fairly small (12 students), and since the students met as a combined group for all other lessons. This would have cast doubt on the experiment's validity. Therefore, the Spring 1999 *Geography 2* students were used as the experimental group, and the control group was comprised of students from the Spring 1996 *Geography 2* class, the last semester in which the unit was run without GIS. Control group students used traditional materials—a table of demographic variables, census tract base maps of the city, and colored pencils for use in creating choropleth maps of the data.

Completed *Hill* projects were scored by a scoring guide in which each criterion had a maximum possible score of 100. The final score was the mean of the five criteria on the left side of the table, and had a maximum possible score of 100 (Table 4.11). I assessed essays based on whether they answered the question assigned, demonstrated understanding, and had original interpretations, logic and transitions in the argument, and required mechanics and style.

Two-sample *t*-tests were run on each criterion. None of the tests showed a significant difference. The highest *t*-statistic was on the grade for the assignment assessed by the teacher ($t=1.4601$), not high enough to be significant. For only one criterion did the experimental group score higher on average than did the control group—for "evidence of content knowledge" (72 vs. 66.67). However, according to the teacher, one-fourth of the control group students referred to their maps as analytical tools to support the arguments in their essays, compared with two-thirds of experimental group students. Small sample sizes (3 cases for the control group and 5 for the experimental group) cast doubt on the generalizability of these data.

Table 4.11. Scoring Guide for *The Hill* Assessment; Maximum Score = 100.

| Assessment Criteria | Score 1; range 0-100 | Score 2; range 0-100 | Score 3; range 0-100 | Score 4; range 0-100 | Score 5; range 0-100 |
|---|---|---|---|---|---|
| **Grade Assessed by Teacher** | Grade, converted to a percentage | | | | |
| **References to Maps as Analytic Tools** | No references | 1 reference | 2 references | 3 references | 4 or more references |
| **Persuasive-ness of Argument** | No real argument or case | Argument is there, but weak | Builds case, but contains gaps | Builds strong case | Builds very strong case |
| **Number of Factors Cited** | No factors | 1-2 factors | 3-4 factors | 5-6 factors | 7 or more factors |
| **Evidence of Content Knowledge** | Below standard; Less than one paragraph | Little evidence; one paragraph | Some evidence; 2 paragraphs | Much evidence; 3 para-graphs | Greatly exceeds standard; 4 or more para-graphs |

***Analysis of Final Course Grades***

Final course grades at Riparian High School were examined to determine if GIS affects overall achievement in a geography course, and whether GIS has a differential effect on low versus high-achieving students. A one-way analysis of variance (ANOVA) with a Scheffe multiple comparison test was conducted for all sections of Grade 9 *Geography 1* (Table 4.12). and found a mean grade of 75.80 (n=51; standard deviation=15.32) for the students using traditional methods, and a mean grade of 77.63 for students using GIS (n=60; standard deviation=13.64).

Students using traditional methods and materials received a “C” grade, on average, and GIS students received a C grade that was a bit higher toward a “B”. A frequency analysis revealed that approximately 17.6% of traditional students received an “A” in the course, while 20% of GIS-using students did. However, the difference between the final grades of the two groups was not significant (t=0.6654; P=.5072) and similar standard deviations indicated that the ranges of scores for the two groups were similar.

Standardized test scores were compared, each accounting for the final course grade of each student. This was accomplished through six paired *t*-tests, considering control and experimental groups as two groups in the school, regardless of class period. “A” students (scoring over 90 on the final course grade) showed more improvement using traditional materials than GIS, but “C” and “D” students showed more improvement using GIS (Table 4.12). Gains were significant for both groups of “A” students, but only “C” and “D” students using GIS posted significant gains. This suggests that GIS might potentially offer more benefits for average and below-average students. Possible reasons will be explored in Chapter 5.

Table 4.12. Results of 6 Paired *t*-tests on Changes on Standardized Tests, Control and Experimental Groups, Based on Final Course Grade (* = significant at ∀=.05; ** = significant at ∀=.10).

| | **Final Course Grade** | | |
|---|---|---|---|
| | D: 60 – 69 | C: 70 – 79 | A: 90 – 100 |
| Experimental Groups (GIS) | t = 2.2863<br>P = 0.0516 ** | t = 3.3874<br>P = 0.0096 * | t = 2.3913<br>P = 0.0233 * |
| Control Groups (no GIS) | t =0.7322<br>P = 0.5170 | t =2.0839<br>P = 0.0613 | t =5.6196<br>P = 0.0000 * |

***Gender Analysis***

An ANOVA with a Scheffe multiple comparison test was conducted for both the control and experimental groups at Riparian High School to find out whether final grades varied with gender. The mean grade for females (79.68; n=60) was significantly higher than that for males (73.39; n=51)(F=5.48; t=-2.3403; P=0.0211). The ANOVA F-test leads a rejection of the null hypothesis that the two groups' population means are equal. Bartlett's test for equal variances ($\chi^2$=0.9816; P=0.322) gives no reason to doubt the equal-variance assumption upon which ANOVA rests. Two-sample *t*-tests showed no significant difference between males and females on the spatial analysis pretest, spatial analysis posttest, standardized pretest, and standardized posttest (Table 4.13). Interestingly, female students performed better than males on the spatial analysis test, while males out-performed females on the standardized test. During the semester, males improved more than females, *losing fewer* points on the spatial test and *gaining more* points on the standardized test.

Table 4.13. Results of 4 Two-sample *t*-tests on Pretests and Posttests, by Gender, Riparian High School.

| **Class Period** | **Test** | **n** | **Mean** | **Standard Deviation** | **Difference** | **t, P values** |
|---|---|---|---|---|---|---|
| Spatial Analysis Pretests | Males | 54 | 12.4630 | 4.3641 | 1.2558 | t=1.7053 P=0.0908 |
| | Females | 64 | 13.7188 | 3.6361 | | |
| Spatial Analysis Posttests | Males | 47 | 9.8936 | 4.6449 | .9708 | t=0.6579 P=0.5120 |
| | Females | 62 | 10.5323 | 5.2844 | | |
| Standardized Pretests | Males | 55 | 65.4088 | 17.0153 | -4.9955 | t=-1.7481 P=0.0830 |
| | Females | 67 | 60.4133 | 14.5469 | | |
| Standardized Posttests | Males | 46 | 72.1083 | 14.2750 | -2.8660 | t=-1.1090 P=0.2701 |
| | Females | 54 | 69.2523 | 11.4718 | | |

The above gender analysis did not consider whether students used GIS during the semester. Therefore, eight two-sample *t*-tests were run, considering whether the student was male or female and whether they were in the control or experimental group (Table 4.14). The only significant difference found was that females who used GIS scored significantly higher on the spatial analysis posttests (t=-2.3105; P=0.0243) than females who used traditional methods and materials. However, using the Bonferroni correction, accounting for multiple *t*-tests, not even this difference would have been significant. Therefore, GIS was not found to make a significant difference in performance by either gender.

Table 4.14. Results of Two-sample *t*-tests on Gender, Pretests, and Posttests, by Control and Experimental Group, Riparian High School (*=significant at ∀=.05).

| ***t*-test Results: Control vs. Experimental Groups** | | | |
|---|---|---|---|
| **Two-sample *t*-test, by the use of GIS** | **Gender** | ***t*-statistic** | **P** |
| Spatial Analysis Pretest | Males | -.7741 | .4424 |
| Spatial Analysis Posttest | Males | 1.4259 | .1608 |
| Standardized Pretest | Males | .7557 | .4532 |
| Standardized Posttest | Males | -.2619 | .7947 |
| Spatial Analysis Pretest | Females | 1.4610 | .1491 |
| Spatial Analysis Posttest | Females | 2.3105 * | .0243 * |
| Standardized Pretest | Females | .0486 | .9614 |
| Standardized Posttest | Females | -.1960 | .8453 |

***Regression Analysis***

Regression analysis was employed to further explore the relationship between the presence of GIS and the difference in pretest versus posttest scores. Regression analysis not only allowed the determination of how strongly related these variables were, but measured the extent of the effect of a change in the independent variables on the dependent variable. In each of the following analyses, the dependent variable was the difference between the pretest versus posttest scores. The presence or absence of GIS acted as a dummy independent variable. Large sample sizes and few variables permitted the reporting of $R^2$ values rather than adjusted $R^2$ values.

The first regression analysis was on the improvement in the spatial analysis test score, controlling for the spatial analysis pretest score, resulting from the presence of GIS (Table 4.15). The pretest score was included because improvement is expected to be lower for those with higher pretest scores. Higher pretest scores leave less room for improvement than lower scores; that is, a student scoring a 19 out of 20 on the spatial analysis pretest cannot improve to the same extent on the posttest that a student scoring a 4 on the pretest can. The dependent variable is the difference between the spatial analysis pretest and the posttest, with pretest score and a dummy variable of the presence (1) or absence (0) of GIS as independent variables.

Table 4.15. Regression Analysis of Spatial Analysis Pretest and Posttest Differences, Controlling for Spatial Analysis Pretest Score, by the Presence of GIS, Riparian High School (*=significant at ∀=.05).

| **Regression Analysis: Difference Between Spatial Analysis Pretest and Posttest** <br> **$R^2$ = .2779 N=100 F(2,97)=18.66** | | | |
|---|---|---|---|
| **Variable** | **Coefficient** | ***t-statistic*** | **P** |
| Spatial Analysis Pretest Score | -.7233 | -5.814 | 0.000* |
| Presence of GIS | 2.2786 | 2.289 | 0.019* |
| Constant | 5.2907 | 2.992 | 0.004* |

These variables *do* affect the difference between the pretest and posttest scores, explaining 27.8% of the variation of the difference score. The resulting equation of the relationships, using the partial regression coefficients that represent the effect of the independent variables on the dependent variable, is:

| | | | | |
|---|---|---|---|---|
| Spatial Analysis Difference Score | = 5.29 | + 2.28 *(Dummy variable GIS or no GIS)* | - | .72 (*pretest score)* |

The *t*-statistics indicate that the pretest's effect on the difference score is significant, and the presence of GIS is also significant. A student scoring a 0 on the pretest can expect to increase to 5.29 on the posttest. The use of GIS increases the difference between the pretest and posttest score by an average of 2.28 points. As suspected, the pretest score is negatively related to improvement (-.723), suggesting that those with lower pretest scores improve more than those with higher pretest scores. Figure 4.1 shows these relationships.

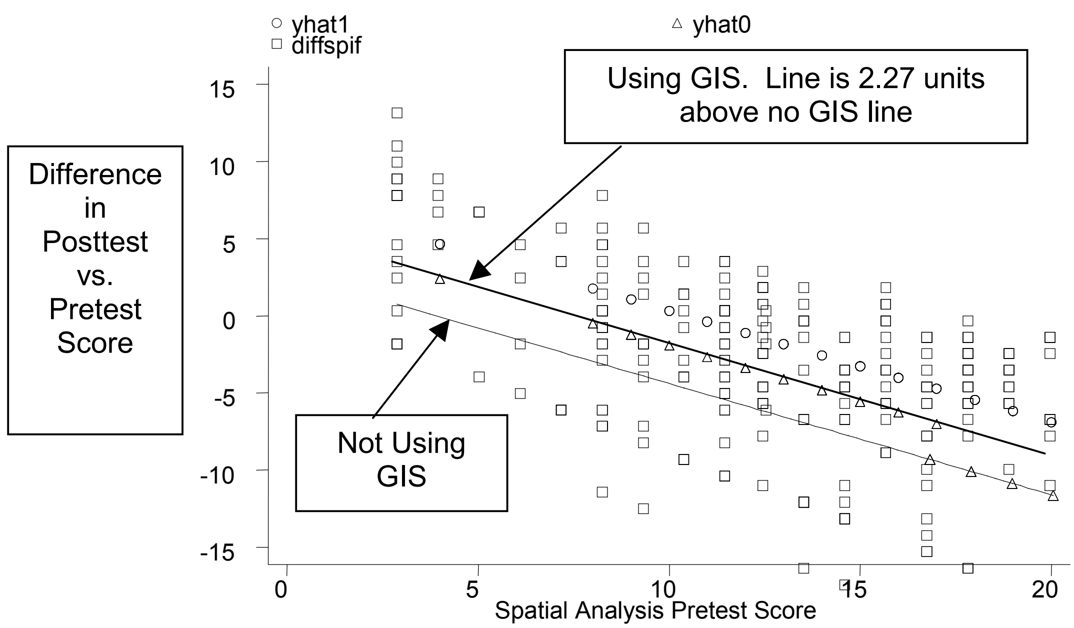


Figure 4.1. Regression Analysis of Improvement of Spatial Analysis Test Scores, Controlling for Pretest Score, by the Use of GIS, Riparian High School.

A regression analysis was also conducted using the difference between the standardized pretest versus posttest score (Table 4.16):

Table 4.16. Regression Analysis of Standardized Pretest and Posttest Differences, Controlling for Standardized Pretest Score, by the Presence of GIS, Riparian High School (*=significant at ∀=.05).

| **Regression Analysis: Difference Between Standardized Pretest and Posttest**<br><br>**$R^2$ = .2817 N=94 F(2,91)=17.85** | | | |
|---|---|---|---|
| **Variable** | **Coefficient** | ***t-statistic*** | **P** |
| Standardized Pretest Score | -.3889 | -5.970 | 0.000* |
| Presence of GIS | -.6597 | -0.343 | 0.732 |
| Constant | 32.1046 | 7.386 | 0.000* |

Unlike the spatial analysis model, the use of GIS in this standardized model does not significantly influence the standardized test scores. The pretest score again negatively affects the difference score, and the model explains 28.2% of the variation of the difference. The *t*-statistics indicate that the pretest's effect on improvement is negative and significant. The GIS coefficient is non-significant at any acceptable probability level in the standardized model, but the GIS coefficient in the spatial analysis model is significant at the .05 level. This suggests that GIS may influence spatial analysis and problem-solving more than traditional locational geography skills.

The above regression models account for the pretest score, but they assume a linear fit to the graph of difference scores by pretest score. Allowing the line that best represents the model to be a curve would help determine if a non-linear relationship exists between performance and GIS. To accomplish this, an independent variable was added to the regression model representing the square of the spatial analysis pretest score. The squared term allows the learning effect to vary depending on the pretest score.

Regressing the difference in spatial analysis scores from pretest to posttest on the spatial analysis pretest score, the square of the pretest score, and the presence of GIS explains 38.4% of the variation of the difference score (Table 4.17):

Table 4.17. Regression Analysis of Standardized Pretest and Posttest Differences, Controlling for Spatial Analysis Pretest Score, by the Presence of GIS, Non-Linear Relationship, Riparian High School (*=significant at ∀=.05).

| **Regression Analysis: Difference Between Spatial Analysis Pretest and Posttest, Non-Linear Model** <br> **$R^2$ = .3839 N=209 F(3, 205)=42.58** | | | |
|---|---|---|---|
| **Variable** | **Coefficient** | ***t-statistic*** | **P** |
| Spatial Analysis Pretest Score | -1.7629 | -4.766 | 0.000* |
| Square of Spatial Analysis Pretest Score | .0407 | 2.570 | 0.011* |
| Presence of GIS | 1.0667 | 1.569 | 0.118 |
| Constant | 13.10553 | 6.433 | 0.000* |

Students using GIS scored 1.067 points better than students using traditional methods. This was not a significant improvement, but it does show that GIS made a positive difference and the difference varied with the pretest score. A similar regression using the difference in *standardized* scores explained nearly 42% of the variation of the difference score, but the coefficient for the presence of GIS was negative and insignificant. Thus, GIS had some influence on performance at Riparian High School, and these influences will be compared with the other two high schools after the other schools are analyzed individually.

## Analysis of Experiments at Hope High School

### *Description of Geography Program and GIS Implementation*

The geography curriculum at Hope consists of *Geography*, a semester-long, elective class for Grade 9 students, and *Advanced Geography*, a semester-long, elective class for Grade 11 and 12 students. The implementation of GIS at Hope High School resulted from a National Geographic Society Education Foundation Grant that this author co-wrote to introduce geographic technology into two Colorado school districts, including the district where Hope High School is located. Ms. Diana Cessna, one of Hope's geography teachers, participated in the grant-sponsored GIS training during 1997 and 1998. At the same time, she implemented GIS in her *Advanced Geography* classes. The year of the experiments was the second year of GIS implementation at the high school. Ms. Cessna has been teaching geography for 10 years, and started teaching at Hope High School in 1995. She has a Bachelor's degree in International Relations and a Master's degree in Special Education Affective Needs. Ms. April Eliot, the teacher who instructed one of the control groups, has been teaching for nine years, including four years of teaching geography, and three years at Hope. She has a background in history, anthropology, and English with a minor in geography. At the time of the experiments, she knew very little about GIS, but soon became interested enough to participate in the school district's National Geographic GIS initiative during the fall semester following the experiments.

### *Description of Experiments*

Experiments conducted at Hope High School took place during Fall semester 1998 and Spring semester 1999 (Table 4.18). The *Pump It-Burn It-Ship It* oil lesson (Appendix A.30) grew out of Mr. Cessna's desire to analyze energy data from the Internet within a GIS. The activity involves the temporal and spatial analysis of oil production, consumption, and reserves from the British Petroleum web site. Experiments were conducted as planned, except that few GIS students were able to complete the oil lesson, and therefore, no comparison could be made between control and experimental groups.

Table 4.18. Experiments Conducted in Hope High School.

| Experi-ment | Class | Grade Level | Lessons | Description | Assess-ment |
|---|---|---|---|---|---|
| **3**:<br><br>Compare C3 Vs E3 | *Advanced Geography* Fall 1998 | 11 and 12 | Pump It-Burn It-Ship It Oil Lesson<br><br>Earthquakes Everyday | Control Group = Period 4 (1240 – 205 pm). Teacher= Cessna<br><br>Experimental Groups = Period 1 (745-910am) and Period 2 (915 – 1040am) Teacher= Cessna | School District standard-ized pretest and posttest |
| **4:**<br><br>Compare C4 vs E4 | *Advanced Geography* Spring 1999 | 11 and 12 | County Demographics<br><br>Earthquakes Everyday<br><br>Pump It-Burn It-Ship It Oil Lesson | Control Group = Period 3 (1045am-12pm) Teacher= Eliot<br><br>Experimental Group = Period 6 (2-315pm) Teacher= Cessna | School District standard-ized pretest and posttest<br><br>Spatial Analysis pretest and posttest |

In Experiment 3, the experimental group was taught two lessons using GIS, while the control group used traditional materials. In Experiment 4, experimental group students were given three GIS-based lessons. Students were given the same spatial analysis test as that given to Riparian High School students. The standardized test was similar—based on the national geography standards, but in this case written by the County Assessment Board and based on national geography standards 1 (using maps) and 4 (physical and human characteristics of places). This test required students to use *Goodes World Atlas* (Appendix A.8).

### *Tests for Group Differences*

To determine if the control and the experimental groups had any pre-existing significant difference before the experiments were constructed, a two-sample *t*-test with equal variances was run on the spatial analysis test and the standardized test. Although the spatial analysis test showed that no significant difference existed (t=-0.9594; P=.3438), the standardized test results showed that the experimental group scored significantly higher than the control group (t=-6.6727; P=0.0000). This difference existed even when *t*-testing experiment 3 and 4 separately. Care therefore needs to be given in assessing these two groups, since they performed differently at the beginning of the semester on one assessment.

### *Assessing Spatial Analysis Tests*

Paired *t*-tests conducted for the spatial analysis tests for the Spring 1999 control group showed an increase in student performance from the pretest to the posttest, but the increase was not significant (t=1.2710; P=.2260). The mean score increased slightly, from 12.78 to 14.29 (out of 20). To provide another control group, four Grade 9 sections of Mr. Asi's geography class at Hope were given the spatial analysis test at the beginning and at the end of the semester. No significant difference was found between the mean pretest and posttest scores (10.1 vs. 10.65). Similarly, paired t-tests conducted for the spatial analysis tests for the Spring 1999 experimental group showed that student performance increased slightly but insignificantly ($t=.1185$; $P=.9073$).

### *Assessing Standardized Tests*

Four paired *t*-tests for the standardized tests for control groups and experimental groups for Experiments 3 and 4 (Fall and Spring semesters) showed significant differences in three cases (Table 4.19). Scores increased in all cases, but students using GIS did not improve significantly more than those using traditional methods.

Table 4.19. Results of Paired *t*-tests on Standardized Tests, Hope High School; Separate Groups (*=significant at ∀=.05).

| Group | Test | n | Mean | Standard Deviation | Difference | t, P values |
|---|---|---|---|---|---|---|
| Experiment 3 Control Group | Pretest | 24 | 66.0000 | 7.7347 | 14.2273 | t=-6.1077 P=0.0000* |
| | Posttest | 24 | 80.2273 | 9.3529 | | |
| Experiment 4 Control Group | Pretest | 22 | 43.1469 | 21.8423 | 37.0629 | t=6.1162 P=0.0000* |
| | Posttest | 22 | 80.2098 | 26.8672 | | |
| Experiment 3 Experimental Group | Pretest | 49 | 75.6939 | 8.5931 | 7.7570 | t= 3.8930 P=0.003* |
| | Posttest | 49 | 83.4508 | 10.0703 | | |
| Experiment 4 Experimental Group | Pretest | 20 | 80.9231 | 14.5815 | 3.6154 | t=1.2154 P=0.2391 |
| | Posttest | 20 | 84.5385 | 6.8525 | | |

The above analysis considers each experiment separately. Combining the two control groups into one control group, and doing the same for the experimental groups allows for more generalizable comparisons in the high school. Paired *t*-tests showed significant differences for both groups (Table 4.20). Scores increased in all cases, but students using GIS did not improve significantly more than those using traditional methods; indeed, the difference was more pronounced for the control group because their pretests were so much lower (t=3.94 for GIS students vs. 7.1034 for control group). GIS students' posttest scores were 3.5 points higher than students who did not use GIS.

Table 4.20. Results of Paired *t*-tests on Standardized Tests, Hope High School; Combined Groups (*=significant at ∀=.05).

| Group | Test | n | Mean | Standard Deviation | Difference | t, P values |
|---|---|---|---|---|---|---|
| Control Group | Pretest | 46 | 55.0702 | 19.6579 | 25.1487 | t=7.1034 P=0.0000* |
| | Posttest | 46 | 80.2189 | 19.5339 | | |
| Experimental Group | Pretest | 69 | 77.2096 | 10.8279 | 6.5565 | t=3.9475 P=0.0002* |
| | Posttest | 69 | 83.7661 | 9.2169 | | |

***Assessing County Social Area Analysis and Earthquake Lessons***

A comparison of Hope student performance *on the actual lessons* would be instructive to assess whether GIS made any difference, and the amount of that difference. The County Social Area Analysis lesson required students to explore spatial and causal relationships by comparing a set of maps containing census housing and demographic variables with a set of educational variables (Appendix A.26 and A.27). Students analyzed such educational variables as mobility rates, dropout rates, and SAT and ACT scores by high school attendance area. For other variables, census tracts and block groups were the geographic units of analysis. After students analyzed whether there were relationships, for example, between the test scores and family structure at the county level, they analyzed similar variables for countries around the world. Experimental groups were provided with data in *ArcView* shape files (digital maps) and database tables, while the control groups' materials were maps and tables plotted from the GIS software.

A two-sample t-test conducted on scores from the County Social Area Analysis lesson showed that students using GIS scored significantly better than for the control group (t=3.9913; P=0.003). GIS students' mean score was 48.4 out of 50, compared to 42 for students not using GIS, showing that students using GIS demonstrated significantly better knowledge and skills.

Once again, students were not able to complete the *Earthquakes Everyday* lesson. I could not evaluate the entire test for comparison purposes, but I evaluated the question asking three noticeable characteristics of the pattern of earthquakes, and scored according to the scoring guide in Table 4.10 (page 190). No significant difference was noted by a two-sample *t*-test between the control and experimental groups in Experiment 3 (t=.6231; P=.5366). However, in Experiment 4, experimental group students performed significantly better (t=-4.0512; P=0.001).

### *Assessing Gender Differences*

Two-sample *t*-tests showed no significant difference between males and females on the spatial analysis pretest, spatial analysis posttest, standardized pretest, and standardized posttest (Table 4.21). Although test results showed less gender difference than for Riparian students, females improved to a greater degree during the semester than did males. A two-sample t-test on student grades on the County demographic lesson also found no significant gender difference, with females scoring two points on average below males (t=1.2087; P=.2345).

Table 4.21. Results of 4 Two-sample *t*-tests on Pretests and Posttests, by Gender, Hope High School.

| Class Period | Test | n | Mean | Standard Deviation | Difference | t, P values |
|---|---|---|---|---|---|---|
| Spatial Analysis Pretests | Males | 17 | 13.2353 | 1.0797 | -.2353 | t=-0.1609 P=0.8732 |
| | Females | 18 | 13.0000 | .9901 | | |
| Spatial Analysis Posttests | Males | 16 | 13.8750 | 3.0957 | .7917 | t=0.8453 P=0.4042 |
| | Females | 18 | 14.6667 | 2.3515 | | |
| Standardized Pretests | Males | 59 | 71.1069 | 17.2673 | -3.8258 | t=-1.1625 P=0.2475 |
| | Females | 55 | 67.2811 | 17.8660 | | |
| Standardized Posttests | Males | 59 | 83.7810 | 8.5550 | -.1653 | t=-0.0945 P=0.9249 |
| | Females | 54 | 83.6156 | 10.2524 | | |

***Regression Analysis***

Regression analysis was used to further explore the relationship between the presence of GIS and the difference in pretest versus posttest scores. A model was created that considered the change in spatial analysis and standardized test scores, controlling for the pretest score, resulting from the presence of GIS. In the regression model below (Table 4.22; Figure 4.2), the dependent variable is the difference between the spatial analysis pretest and the posttest, with pretest score and a dummy variable of the presence (1) or absence (0) of GIS as independent variables.

These variables do affect the difference between the pretest and posttest scores, explaining 62.2% of the variation of the difference score. However, GIS does not significantly influence the difference in test scores.

Table 4.22. Regression Analysis of Spatial Analysis Pretest and Posttest Differences, Controlling for Spatial Analysis Pretest Score, by the Presence of GIS, Hope High School (*=significant at ∀=.05).

| Regression Analysis: Difference Between Spatial Analysis Pretest and Posttest<br><br>$R^2$ = .6228 N=29 F(2, 26)=21.47 | | | |
|---|---|---|---|
| **Variable** | **Coefficient** | ***t-statistic*** | **P** |
| Spatial Analysis Pretest Score | -.8065 | -6.417 | 0.000* |
| Presence of GIS | -.3335 | -0.319 | 0.752 |
| Constant | 11.8122 | 6.672 | 0.000* |

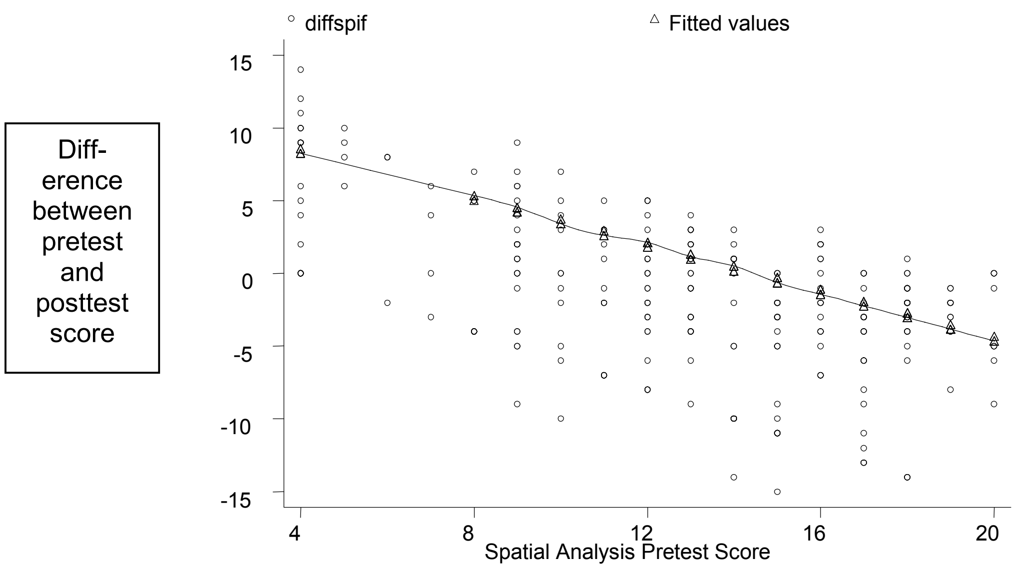


Figure 4.2. Plot Showing Regression Analysis of Spatial Analysis Pretest and Posttest Differences (y-axis), Controlling for Spatial Analysis Pretest Score (x-axis), by the Presence of GIS, Hope High School.

A regression analysis on the difference between standardized pretest versus posttest score found no significant difference (Table 4.23).

Table 4.23. Regression Analysis of Standardized Pretest and Posttest Differences, Controlling for Standardized Pretest Score, by the Presence of GIS, Hope High School (*=significant at ∀=.05).

| Regression Analysis: Difference Between Standardized Pretest and Posttest<br><br>$R^2$ = .5442 N=115 F(2, 112)=66.87 | | | |
|---|---|---|---|
| **Variable** | **Coefficient** | ***t-statistic*** | **P** |
| Standardized Pretest Score | -.8143 | -9.237 | 0.000* |
| Presence of GIS | -.5643 | -0.171 | 0.865 |
| Constant | 69.9918 | 13.263 | 0.000* |

The pretest score again affects the difference score, and the model explains 54.4% of the variation of the difference. The *t*-statistics indicate that the pretest's effect on improvement is negative and significant, but the effect of GIS is insignificant. We have more confidence with this model because the sample size is approximately four times greater than for the spatial analysis test, and the results are similar.

Next, a regression model that helps investigate if a non-linear relationship exists between performance and GIS was created, using the square of the spatial analysis pretest score (Table 4.24).

The pretest score's effect on improvement is negative and significant. Also, students using GIS performed 1.34 points *worse* than students using other materials, although the amount was insignificant. Overall, neither this test nor the standardized nonlinear regression model (Table 4.25) found GIS to be significant in terms of helping students learn.

Table 4.24. Regression Analysis of Spatial Analysis Pretest and Posttest Differences, Controlling for Spatial Analysis Pretest Score, by the Presence of GIS, Non-Linear Relationship, Hope High School (*=significant at ∀=.05).

| Regression Analysis: Difference Between Spatial Analysis Pretest and Posttest, Non-Linear Model<br><br>$R^2$ = .7255 N=29 F(3, 25)=22.03 | | | |
|---|---|---|---|
| **Variable** | **Coefficient** | ***t-statistic*** | **P** |
| Spatial Analysis Pretest Score | -2.3157 | -4.583 | 0.000* |
| Square of Spatial Analysis Pretest Score | .0651 | 3.059 | 0.005* |
| Presence of GIS | -1.3414 | -1.387 | 0.178 |
| Constant | 19.7219 | 6.553 | 0.000* |

The use of GIS does not significantly influence the difference between the standardized pretest and posttest scores, perhaps because it may teach spatial reasoning, rather than ability to complete standardized tests. Students using GIS scored two points *worse* than students using traditional methods, although the difference was insignificant. GIS did have some influence on performance at Hope High School, particularly in the quality of student work on the lessons themselves.

Table 4.25. Regression Analysis of Standardized Pretest and Posttest Differences, Controlling for Standardized Pretest Score, by the Presence of GIS, Non-Linear Relationship, Hope High School (*=significant at ∀=.05).

| Regression Analysis: Difference Between Standardized Pretest and Posttest, Non-Linear Model<br><br>$R^2$ = .3908 N=209 F(3, 205)=43.84 | | | |
|---|---|---|---|
| **Variable** | **Coefficient** | ***t-statistic*** | **P** |
| Standardized Pretest Score | -.6627 | -3.145 | 0.002* |
| Square of Standardized Pretest Score | .0005 | 0.273 | 0.785 |
| Presence of GIS | -2.040 | -1.009 | 0.314 |
| Constant | 53.4946 | 8.635 | 0.000* |

## Analysis of Experiments at Prairie Vista High School

### *Description of Geography Program and GIS Implementation*

The geography curriculum at Prairie Vista High School consists of an elective *World Regional Geography* and two courses that are part of the International Baccalaureate (IB) Program. These courses include a required Pre-IB *World Regional Geography* for IB Grade 9 students and an elective IB *Geography* for Grade 11 and 12 students. For students not in the IB program, the school has no requirement to take a geography course, although most students take the non-IB *World Regional Geography* course in Grade 10.

The International Baccalaureate (IB) program is from the International Baccalaureate Organization (IBO), a nonprofit educational foundation based in Switzerland, which, among its other programs, offers the Diploma Program for students in their final two years of secondary school. Over 1,000 schools are members of the IBO in 100 countries around the world. The IBO grew out of international efforts during the 1960s to establish a common curriculum and university entry credential for geographically mobile students, and a hope that a shared academic experience emphasizing critical thinking and exposure to a variety of viewpoints would foster tolerance and inter-cultural understanding among young people. Concentration on the last two years of secondary school sought to build a comprehensive curriculum that could be administered in any country and recognized by universities in every country. The geography program falls under the IBO's "Individuals and Societies" curriculum. Students must apply to be in the IB program, whose courses are recognized as being quite rigorous. The IB Grade 11-12 *Geography* syllabus looks similar to that of most upper-level *college-level* course outlines.

Geography teacher Richard Clark has a strong background in the discipline, beginning with a bachelor's degree in geography, master's degree in education, and work as a cartographer for the Defense Mapping Agency and regional planning agencies. He had been interested in GIS for many years before working with this author. A National Geographic Society Education Foundation grant proposal he co-wrote with this author was funded for $6884.00, making the GIS vision a reality. Mr. Clark has taught geography since 1993, and began teaching at Prairie Vista in 1995. He also taught U.S. History for four years at the school.

***Description of Experiments***

Experiments conducted at Prairie Vista High School took place during Spring semester 1999 (Table 4.26). Mr. Clark taught all classes.

The semester schedule prohibited students from taking the standardized test, and spatial analysis posttests were not given to Grade 9 students, so neither could be used. The Africa lesson could not be used because of computer lab problems.

Table 4.26. Experiments Conducted in Prairie Vista High School.

| Experi-ment | Class | Grade Level | Lessons | Description | Assessment |
|---|---|---|---|---|---|
| **5:**<br><br>Compare C5 vs. E5 | *Pre-IB World Geography* | 9 | Earthquakes Everyday | Control Group = Period 2 (745-833am).<br><br>Experimental Group = Period 7 (1238-130pm)<br><br>Experimental Group = Period 8 (138 – 230pm) | Standardized test<br><br>Spatial Analysis pretest and posttest |
| **6:**<br><br>Compare C6 vs. E6 | *Advanced Geography* Spring 1999 | 11 and 12 | County Social Area Analysis and Field Experience | Control Group = Period 5 (1038am-1130am)<br><br>Experimental Group = Period 4 (945-1033am) | Standardized test<br><br>Spatial Analysis pretest and posttest |

***Tests for Group Differences***

A two-sample *t*-test with equal variances on the spatial analysis test for Advanced Geography students showed no significant difference between groups (t=-0.3730; P=0.7126). Confidence with subsequent tests comparing the control and experimental groups is high, since there is a good chance that both groups are drawn from the same population.

### *Assessing Spatial Analysis Tests*

To assess the difference between pretest and posttest scores, paired *t*-tests were conducted for each advanced class. Scores for both groups declined, and the experimental group's scores declined significantly (Table 4.27).

Table 4.27. Results of Paired *t*-tests, Prairie Vista High School
(*=significant at ∀=.05).

| Class Period | Test | n | Mean | Standard Deviation | Difference | t, P values |
|---|---|---|---|---|---|---|
| 4 Experi-mental | Pretest | 15 | 17.8667 | 2.3258 | -1.4667 | t=-2.3227 P=0.0358* |
| | Posttest | 15 | 16.4000 | 1.7647 | | |
| 5 Control | Pretest | 8 | 17.5000 | 2.2039 | -0.8750 | t= -1.1784 P=0.2771 |
| | Posttest | 8 | 16.6250 | 1.0607 | | |

As was the case with Riparian High School, these students had a disincentive to thoughtfully complete the test at the end of the semester. Prairie Vista's students scored higher on this test than the other two high schools, possibly reflecting the fact that they are part of the IB program. Low sample sizes, however, require further analyses to be taken, such as assessing the actual lessons the students had completed.

### *Assessing County Social Area Analysis and Field Work Lesson*

The County Social Area Analysis and Field Work lesson required students to explore spatial patterns and relationships in a county on the other side of the Denver metropolitan area, and then comparing their results to field work (Appendix A.28 and A.29). This county was chosen because students would have less direct experience and stereotypes about it than the county in which their school was located, and because the county contains a wider variety of landforms and demographic characteristics. This was the same county that Hope's students analyzed. Students analyzed such variables as median income, housing value and tenure (owning versus renting), and ethnicity by census tract and block group. Experimental group students were provided with data in *ArcView* shape files (digital maps) and database tables, while the control groups' materials were maps and tables plotted from the GIS software. A combined assembly of students from both groups then took a guided field trip of the county, answering a series of essay questions en route about the characteristics of each neighborhood, and whether these observed characteristics matched their expectations.

A two-sample t-test conducted on scores from the County Social Area Analysis and Field Work lesson showed no significant difference between students using GIS and those using paper maps and tables ($t=-1.5666$; $P=0.1315$). Out of 40 points, the mean experimental group score was 1.6 points *below* that of the control group (35.4 vs. 37). Therefore, GIS did not have any demonstrable effect on scores.

Most students were not able to complete the *Earthquakes Everyday* lesson, analogous to the situation at the other schools, nor was the question on the pattern of earthquakes able to be evaluated.

### *Regression Analysis*

Regression analysis considered the change in spatial analysis test scores, controlling for the pretest score, resulting from the presence of GIS. In the regression model below (Table 4.28), the dependent variable is the difference between the spatial analysis pretest and the posttest, with pretest score and a dummy variable of the presence (1) or absence (0) of GIS as independent variables.

Table 4.28. Regression Analysis of Spatial Analysis Pretest and Posttest Differences, Controlling for Spatial Analysis Pretest Score, by the Presence of GIS, Prairie Vista High School (*=significant at ∀=.05).

| **Regression Analysis: Difference Between Spatial Analysis Pretest and Posttest** <br> **$R^2$ = .6004 N=23 F(2, 20)=15.02** | | | |
|---|---|---|---|
| **Variable** | **Coefficient** | ***t-statistic*** | **P** |
| Spatial Analysis Pretest Score | -.7877 | -5.410 | 0.000* |
| Presence of GIS | -.3029 | -0.452 | 0.656 |
| Constant | 12.9092 | 4.956 | 0.000* |

These variables do affect the difference between the pretest and posttest scores, explaining 60% of the variation of the difference score. However, GIS does not significantly influence the difference in test scores.

Next, I created a regression model to investigate whether a non-linear relationship exists between performance and GIS, using the square of the spatial analysis pretest score. In this model, the pretest score's effect on improvement is insignificant, and GIS does not appear to influence the difference in scores. Therefore, none of these tests showed that GIS made a difference in terms of student learning at Prairie Vista High School.

## Combined Analysis of All Case Study Schools

Thus far, this chapter has analyzed student performance in each of three high schools separately. Comparing control and experimental groups that are combined from all three schools will aid in understanding the overall effect of GIS on high school students' learning.

### *Tests for Group Differences*

Before a comparison between control and experimental groups can be confidently made, two-sample *t*-tests compared pretest scores among students that did not use GIS across all three schools. No significant difference was found (t=-1.1788; P=.2400) among the 184 students tested, but a significant difference was found using the same procedure with the standardized test (t=5.3755; P=.0000). This was largely a result of the contribution from Hope High School.

### *Examining Pretest and Posttest Scores*

Two of the four two-sample *t*-tests conducted to assess the difference that GIS made showed significant differences between the control and experimental groups (Table 4.29). First, students using GIS did better at spatial analysis at the end of the semester than their counterparts who used traditional methods (P=.0150). However, because spatial analysis scores declined from pretest to posttest for both groups, it is more accurate to say that scores *declined significantly less* for GIS students than for non-GIS students. Second, standardized pretest scores were significantly higher for GIS-using students, largely because of the influence of Hope High School. Because of this precondition, other experiments must be conducted to fully assess if GIS made a difference in test scores. However, both groups did improve on the standardized test, and the GIS students scored slightly, though insignificantly, higher at the end of the semester.

Table 4.29. Results of Two-sample *t*-tests on Pretest and Posttest Score, All High Schools (*=significant at ∀=.05).

| Test | Group | n | Mean | Standard Deviation | Difference | t, P values |
|---|---|---|---|---|---|---|
| Spatial Analysis Pretest | Control | 79 | 13.2531 | 3.9854 | .7373 | t=1.1788 P=0.2400 |
| | Experimental | 105 | 13.9905 | 4.3534 | | |
| Spatial Analysis Posttest | Control | 77 | 10.9351 | 5.3171 | 1.8391 | t=2.4583 P=0.0150* |
| | Experimental | 93 | 12.7742 | 4.4383 | | |
| Standardized Pretest | Control | 99 | 58.7309 | 19.4503 | 11.6370 | t=5.3755 P=0.0000* |
| | Experimental | 140 | 70.3679 | 14.0234 | | |
| Standardized Posttest | Control | 96 | 75.6041 | 16.9811 | 2.2217 | t=1.1094 P=0.2685 |
| | Experimental | 123 | 77.8258 | 12.6528 | | |

Four paired *t*-tests on spatial analysis and standardized tests were run to assess the difference that GIS made on student learning *within the same group over the semester* as measured by pretest versus posttest scores (Table 4.30). Although significant differences (even after applying a Bonferroni correction) resulted for each test, GIS did not appear to make any difference in scores. Spatial analysis scores for both the control and the experimental groups declined significantly between the beginning and the end of the semester. Scores for the group using GIS declined slightly *more* than scores for those using traditional methods. Standardized scores showed the opposite trend, increasing for both groups, and increasing slightly more for the control group. Standardized scores were, however, higher for the experimental group.

Table 4.30. Results of Paired *t*-tests on Spatial Analysis and Standardized Tests, All Schools (*=significant at ∀=.05).

| Group | Test | n | Mean | Standard Deviation | Difference | t, P values |
|---|---|---|---|---|---|---|
| Spatial Analysis Control Group | Pretest | 70 | 13.5429 | 4.0600 | -2.6714 | t=-3.8550 P=0.0003* |
| | Posttest | 70 | 10.8714 | 5.2913 | | |
| Spatial Analysis Experimental Group | Pretest | 82 | 14.5854 | 3.9408 | 1.7683 | t=-3.4783 P=0.0008* |
| | Posttest | 82 | 12.8171 | 4.4085 | | |
| Standardized Control Group | Pretest | 94 | 59.3459 | 18.6666 | 16.1005 | t=7.5767 P=0.0000* |
| | Posttest | 94 | 75.4464 | 17.1253 | | |
| Standardized Experimental Group | Pretest | 115 | 71.4939 | 13.4778 | 6.7231 | t=5.7704 P=0.0000* |
| | Posttest | 115 | 78.2170 | 12.7683 | | |

### *Assessing Gender Differences*

Only one significant difference was discovered from six two-sample t-tests conducted on pretest scores, posttest scores, and difference scores that grouped all three high schools together (Table 4.31).

Table 4.31. Results of Two-sample *t*-tests on Pretest and Posttest Score, Gender Differences, All High Schools (*=significant at ∀=.05).

| Test | Group Performing Better or Improving More | *t-statistic* | P |
|---|---|---|---|
| Spatial Analysis Pretest | Females | .9362 | .3505 |
| Spatial Analysis Posttest | Females | .2526 | .8009 |
| Standardized Pretest | Males | 2.2082 | .0282* |
| Standardized Posttest | Males | 1.2684 | .2061 |
| Difference in Spatial Analysis Pretest vs. Posttest Scores | Males[3] | .4649 | .6427 |
| Difference in Standardized Pretest vs. Posttest Scores | Females | .5998 | .5493 |

### *Analysis of Earthquake Lesson*

[3] Both groups declined in scores. Since females declined more than males, males performed better.

The only lesson that could be analyzed across all three high schools was the spatial pattern question in the *Earthquakes Everyday* unit. Even though the experimental group performed better (mean=8.15 vs. 7.83), the difference was not significant (t=.6091; P=.5432).

### *Regression Analysis*

*Examining Pretest and Posttest Scores*

Analyzing the difference in the means with *t*-tests is instructive, but hides the direction and magnitude of factors that may influence test scores. Regression analysis across all three schools was conducted to explore the relationship between the presence of GIS and learning, as measured by the test scores. Again, large sample sizes and few variables permitted the reporting of $R^2$ values rather than adjusted $R^2$ values.

The first regression analysis was on the effect of GIS on the improvement in the spatial analysis test score, controlling for the spatial analysis pretest score. In the regression table below, the dependent variable is the difference between the spatial analysis pretest and the posttest, with pretest score and a dummy variable of the presence (1) or absence (0) of GIS as independent variables (Table 4.32).

Table 4.32. Regression Analysis of Difference Between Spatial Analysis Pretest and Posttest, Controlling for Spatial Analysis Pretest Score, by the Presence of GIS, All Schools (*=significant at ∀=.05).

| Regression Analysis: Difference Between Spatial Analysis Pretest and Posttest<br><br>$R^2$ = .2272 N=152 F(2, 149)=21.90 | | | |
|---|---|---|---|
| **Variable** | **Coefficient** | ***t-statistic*** | **P** |
| Spatial Analysis Pretest Score | -.6106 | -6.507 | 0.000 |
| Presence of GIS | 1.5397 | 2.043 | 0.043* |
| Constant | 5.5981 | 4.044 | 0.000* |

The positive “presence of GIS” coefficient means that GIS students did significantly better by 1.54 points than non-GIS students. Again, those with lower pretest scores improved the most. A similar regression on the standardized scores resulted in a model that explained more of the pretest-posttest variation ($R^2$ =.3906) but GIS had no significant effect on the scores (coefficient=-2.00; t=-.994; P=.322).

Next, allowing the line that best represents the model to be a non-linear relationship is possible by including an independent variable that represents the square of the spatial analysis pretest score. This model explains 29.3% of the variation of the difference score (Table 4.33; Figure 4.3). The presence of GIS was significant at P=.10. GIS has more influence on students who performed poorly on the spatial analysis tests at the beginning of the semester than students who performed well.

Table 4.33. Regression Analysis of Difference Between Spatial Analysis Pretest and Posttest, Controlling for Spatial Analysis Pretest Score, by the Presence of GIS, Non-Linear Model, All Schools (*=significant at ∀=.05; ** at ∀=.10).

| **Regression Analysis: Difference Between Spatial Analysis Pretest and Posttest, Non-Linear Model** <br> **$R^2$ = .2926 N=152 F(3, 148)=20.40** | | | |
|---|---|---|---|
| **Variable** | **Coefficient** | ***t-statistic*** | **P** |
| Spatial Analysis Pretest Score | -2.2616 | -4.967 | 0.000 * |
| Square of Spatial Analysis Pretest Score | .0663 | 3.699 | 0.000 * |
| Presence of GIS | 1.3769 | 1.899 | 0.059 ** |
| Constant | 14.7194 | 5.254 | 0.000 * |

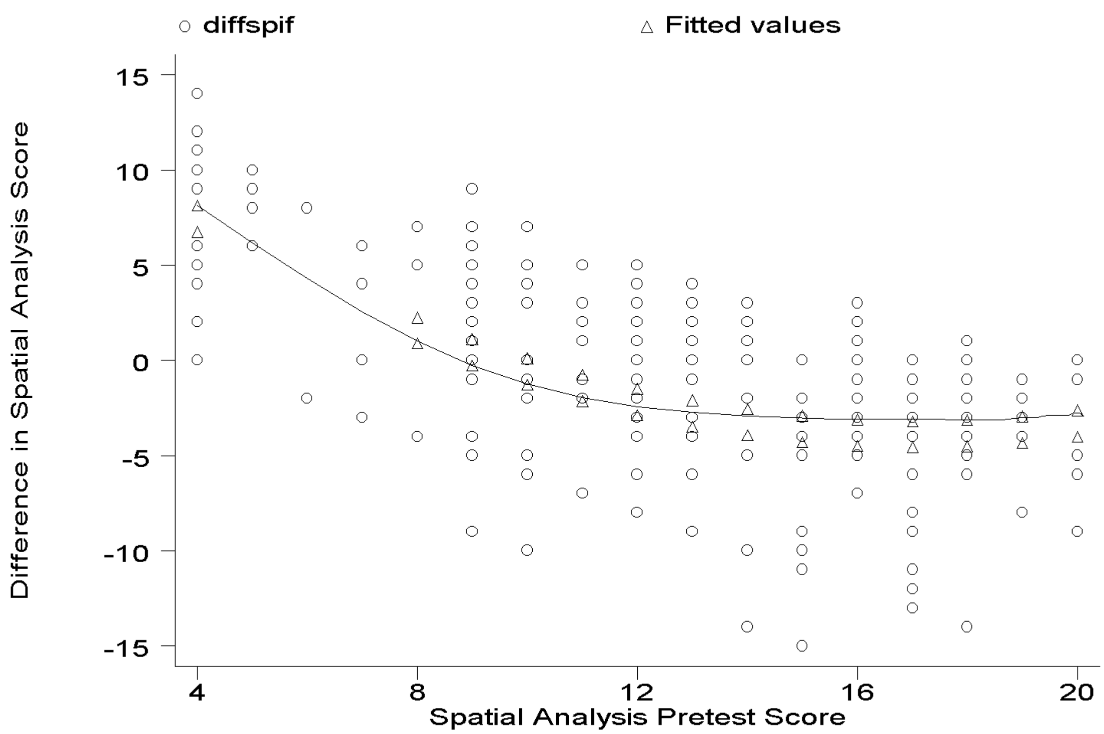


Figure 4.3. Regression Analysis of Spatial Analysis Pretest and Posttest Differences, Controlling for Spatial Analysis Pretest Score, by the Presence of GIS, Non-Linear Relationship, All Schools.

The same non-linear regression on the standardized scores resulted in a model that explained more of the pretest-posttest variation ($R^2$ =.3908) but GIS had an insignificant effect on the scores (coefficient=-2.040; t=-.1.009; P=.314). The low t-statistic for the square of the standardized pretest score (0.273) means that the relationship was nearly linear, and not variable by pretest score (Figure 4.4).

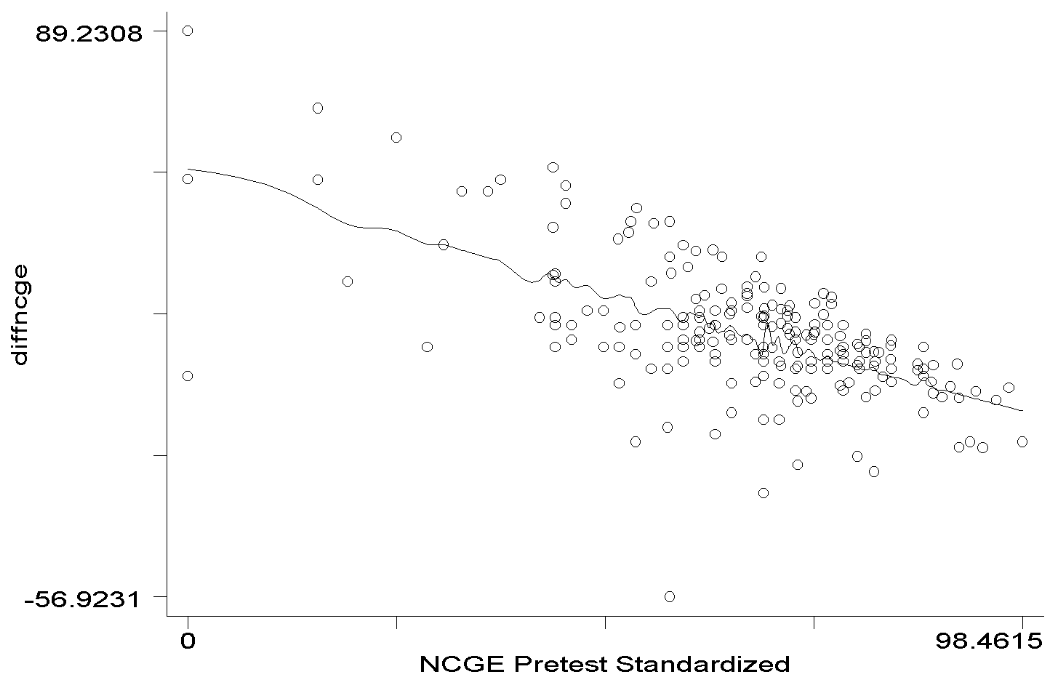


Figure 4.4. Regression Analysis of Standardized Pretest and Posttest Differences, Controlling for Spatial Analysis Pretest Score, by the Presence of GIS, Non-Linear Relationship, All Schools.

*Gender Analysis*

A regression run on the improvement in the spatial analysis score, controlling for the spatial analysis pretest score and considering whether GIS and gender made a difference, showed no gender effect (Table 4.34).

Table 4.34. Regression Analysis of Difference Between Spatial Analysis Pretest and Posttest Scores, Controlling for Spatial Analysis Pretest Score, Considering GIS and Gender; All Schools (*=significant at ∀=.05).

| Regression Analysis: Linear Model: Spatial Analysis-GIS-Gender $R^2$ = .2327 N=146 F(3, 146)=14.76 | | | |
|---|---|---|---|
| **Variable** | **Coefficient** | ***t-statistic*** | **P** |
| Spatial Analysis Pretest Score | -.6126 | -6.498 | 0.000* |
| Presence of GIS | 1.6632 | 2.184 | 0.031* |
| Gender | -.2179 | -0.288 | 0.774 |
| Constant | 5.6731 | 3.934 | 0.000* |

Multiplying gender by the presence of GIS yielded an interaction term that was used to test for differential GIS effects based on gender. It was used in regression models for the spatial analysis and the standardized tests, both with and without the square term of the pretest score. No gender effects were found in either model.

*Teachers' Experience with GIS*

The three schools represented a *range* in the amount of time that GIS has been in use in each school, from first year of implementation (Prairie Vista), second year (Hope), and third year (Riparian). One might expect that as the teachers' experience with GIS tools grows, students perform better. However, analyzing the tests performed on the data in this chapter found no significant increase in scores by students in Riparian High school over scores in Prairie Vista. A variable that captures teachers' experience in GIS would have to be created for further empirical testing. Because each teacher was more experienced with traditional methods, this might have diminished the potential effect of GIS.

## Summary

Eighty-seven tests were conducted on data obtained from six experiments conducted in three high schools. The effectiveness of GIS on student performance using standardized and spatial analysis tests showed mixed results in each school (Tables 4.35, 4.36, and 4.37) and while considering all schools together (Table 4.38). Spatial analysis test scores either did not change or declined between the beginning and end of the semester. Declining student performance suggests inadequacies with the spatial analysis test and a disincentive for students to thoughtfully complete it at the end of the semester. If a teacher uses GIS, he or she is not able to spend as much time on “testable” content that would appear on a standardized test. GIS did not typically appear to affect the stagnant or downward trend in spatial analysis scores. However, linear and non-linear regression models considering all schools showed that GIS did make a difference in the relationship between GIS and the difference in test scores from the beginning to the end of the semester. Standardized test scores showed improvement over the course of the semester for most students, but again, GIS did not appear to affect this improvement. Tests on GIS on final course grades suggest that average and below-average students improve more with GIS than above-average students.

GIS did have a significant effect on student performance on the lessons themselves. In four out of nine tests, students using GIS scored significantly higher than their counterparts who were using traditional methods, and demonstrated a better ability to synthesize, identify, and describe reasons for human and physical patterns (Table 4.39). GIS seems to foster both analytical thinking and synthetic thinking. Students broke apart a problem or issue into workable pieces but could also put different pieces of information together. GIS appears to improve learning of geographic content, not just skills. Furthermore, GIS appeared to not only foster higher-order analytical and synthetic thinking, but it also increased students' knowledge of absolute and relative locations of places—such as countries, rivers, and cities—across the globe.

Table 4.35. Summary of Experiment Results from Riparian High School.

| Test | Group Tested | x= significant at P=.05 | Direction |
|---|---|---|---|
| Pre-vs posttest, spatial analysis | Control # 1 | --- | --- |
| Pre-vs-posttest, spatial analysis | Control # 2 | x | Decline |
| Pre-vs-posttest, spatial analysis | Experimental # 1 | --- | --- |
| Pre-vs-posttest; spatial analysis | Experimental # 2 | x | Decline |
| Pre-vs-posttest; spatial analysis | Experimental # 3 | x | Decline |
| Pre-vs-posttest; spatial analysis | All Control | x | Decline |
| Pre-vs-posttest; spatial analysis | All Experimental | x | Decline |
| Pre-vs-posttest; standardized | Control # 1 | x | Improvement |
| Pre-vs-posttest; standardized | Control # 2 | x | Improvement |
| Pre-vs-posttest; standardized | Experimental # 1 | x | Improvement |
| Pre-vs-posttest; standardized | Experimental # 2 | x | Improvement |
| Pre-vs-posttest; standardized | All Control | x | Improvement |
| Pre-vs-posttest; standardized | All Experimental | x | Improvement |
| Final Grades | Control vs Experi. | --- | --- |
| Grades; "D" students Standardized | All Control | --- | --- |
| Grades; "C" students; standardized | All Control | --- | --- |
| Grades; "A" students; standardized | All Control | x | Improvement |
| Grades; "D" students; standardized | All Experimental | x (P=.10) | Improvement |
| Grades; "C" students; standardized | All Experimental | x | Improvement |
| Grades; "A" students; standardized | All Experimental | x | Improvement |
| Regression on Spatial Analysis Difference Score | All | x | GIS signif. effect |
| Regression on Standardized Difference Score | All | --- | --- |
| Regression on Spatial Analysis Difference Score; Non-Linear Model | All | --- | --- |
| Regression on Standardized Difference Score; Non-Linear Model | All | --- | --- |

Table 4.36. Summary of Experiment Results from Hope High School.

| Test | Group Tested | x= significant at P=.05 | Description |
|---|---|---|---|
| Pre-vs posttest; spatial analysis | Control | --- | --- |
| Pre-vs-posttest; spatial analysis | Experimental | --- | --- |
| Pre-vs-posttest; standardized; Fall semester | Control | x | Improvement |
| Pre-vs-posttest; standardized; Fall semester | Experimental | x | Improvement |
| Pre-vs-posttest; standardized; Spring semester | Control | x | Improvement |
| Pre-vs-posttest; standardized; Spring semester | Experimental | --- | --- |
| Pre-vs-posttest; standardized | All Control | x | Improvement |
| Pre-vs-posttest; standardized | All Experimental | x | Improvement |
| Regression on Spatial Analysis Difference Score | All | --- | --- |
| Regression on Standardized Difference Score | All | --- | --- |
| Regression on Spatial Analysis Difference Score; Non-Linear Model | All | --- | --- |
| Regression on Standardized Difference Score; Non-Linear Model | All | --- | --- |

Table 4.37. Summary of Experiment Results from Prairie Vista High School.

| Test | Group Tested | x= significant at P=.05 | Description |
|---|---|---|---|
| Pre-vs-posttest; spatial analysis | Control | --- | --- |
| Pre-vs-posttest; spatial analysis | Experimental | x | Decline |
| Regression; pre-vs-posttest; spatial analysis | All | --- | --- |
| Regression; pre-vs-posttest; spatial analysis; Non-Linear model | All | --- | --- |

Table 4.38. Summary of Experimental Results Considering All Schools.

| Test | Group Tested | x= significant at P=.05 | Description |
| --- | --- | --- | --- |
| Spatial analysis pretest | Control vs Experimental | --- | --- |
| Spatial analysis posttest | Control vs Experimental | x | Control improved more |
| Standardized pretest | Control vs Experimental | x | Control improved more |
| Standardized posttest | Control vs Experimental | --- | --- |
| Spatial analysis; pretest vs posttest | Control | x | Decline |
| Spatial analysis; pretest vs posttest | Experimental | x | Decline |
| Standardized; pretest vs posttest | Control | x | Improvement |
| Standardized; pretest vs posttest | Experimental | x | Improvement |
| Regression on Spatial Analysis Difference Score | All | x | GIS signif effect |
| Regression on Standardized Difference Score | All | --- | --- |
| Regression on Spatial Analysis Difference Score; Non-Linear Model | All | x (P=.10) | GIS signif effect |
| Regression on Standardized Difference Score; Non-Linear Model | All | --- | --- |

Table 4.39. Summary of Results from Assessed Lessons, All High Schools.

| Test | School | x= significant at P=.05 | Description |
|---|---|---|---|
| Africa Lesson 1; Control vs Experimental | Riparian | x (P=.10) | Experimental was Higher |
| Africa Lesson 2; Control vs Experimental | Riparian | x | Experimental was Higher |
| Earthquake Pattern Question; Control vs Experimental | Riparian | --- | --- |
| The Hill Neighborhood Analysis; Control vs. Experimental | Riparian | --- | --- |
| County Social Area Analysis; Control vs. Experimental | Hope | x | Experimental was Higher |
| Earthquake Pattern Question; Control vs. Experimental; Semester 1 | Hope | --- | --- |
| Earthquake Pattern Question; Control vs. Experimental; Semester 2 | Hope | x | Experimental was Higher |
| County Demographic and Field Work Exercise; Control vs Experimental | Prairie Vista | --- | --- |
| Earthquake Pattern Question; Control vs Experimental | All | --- | --- |

A total of 26 tests were conducted on the effect of GIS on performance by gender. Of these, only four were found to be significant. At Riparian, female students' final grades were significantly higher than male students', and females using GIS scored higher on the standardized test than females not using GIS. Considering all schools, males did significantly better on standardized pretests. A regression analysis on gender considering all schools found GIS to have a significant effect on the difference between spatial analysis pretest and posttest. Far more often, however, neither gender was significantly different than the other in content knowledge and skills necessary to complete the lessons, spatial analysis tests, or standardized tests. Furthermore, evidence showed that, for the most part, neither gender benefits from GIS more than the other.

The independent variables were part, but most likely not all, of the explanation for the differences in the dependent variables. Although the background of teachers was considered, it was not tested, and the background of students in these courses was not obtained. While an attempt was made to avoid designating the control groups to the same time period in each school, performance is affected by the time of day and the day of the week in which students take tests. The three schools were at three different stages in implementing GIS in the curriculum during the academic year when the experiments took place. Riparian High School was in its third year, Hope was in its second year, and Prairie Vista was in its first year. None of the students had used GIS before the year of the experiments, so the effect of this variable lay solely with the teachers. Finally, because all teachers had more experience using traditional methods than GIS methods, this could have served to lessen the effectiveness of GIS.

It is extremely difficult to isolate the effect of inquiry-oriented approaches such as GIS. Studies such as Hill et al. (1994) that test the “geographic perspective” often show no significant difference because of this difficulty. Geography influences knowledge in other subjects, and other subjects enhance geographic knowledge (for example, knowing mathematics often helps with map scale). Spatial ability influences performance on geography tests. Geographic reasoning weaves together landscape, maps, hypotheses, processes, and models to create an argument or case (Gregg and Leinhardt 1994). Problems in GIS and problems in geography do not usually have just one right answer. Some geographic reasoning is analytical and sequential, but some involves images and requires holistic spatial thinking. As Downs (1994a) pointed out, the process of geography learning takes place both in formal and informal contexts, in and out of school.

These experiments attempted to find out whether students are learning strategies for geographic inquiry, or whether they are merely learning to use a powerful computer tool without thinking about the underlying issues and problems. There is some evidence for the effectiveness of GIS, but the evidence is spotty. Standardized and spatial analysis tests did not fully assess the skills that the research literature revealed that students are gaining with GIS. The standardized test was a typical geography test involving geographic facts and some spatial reasoning. Although the spatial analysis test was created to provide a tool that could more completely assess the skills covered by the GIS lessons, it too was insufficient. Students may have gained knowledge and skills that were not assessed by either test, such as being able to access and use information from various formats, manage a database, use analytical software, and thinking about and solving a problem. Thus, the experiments do not tell the whole story of the effect of GIS in education. To more fully understand the effect, case studies are needed.

Having analyzed the effectiveness of GIS in secondary education at three high schools, the next chapter addresses its implementation and effectiveness at the same schools using a qualitative, case study methodology.